

应用层读书笔记

17211052 王宝源

一、文献信息

论文题目：从图片中的土壤信息和特性推断土壤的病虫害情况。

论文作者：Bruno Ssekiwera & Claire Babirye

发表途径：乌干达技术管理大学

二、问题意义

众所周知，乌干达出口总额的 80% 是农产品，主要是咖啡、茶、棉花等，其中以咖啡为主，占出口总额的 22%。然而，在上一个财政年度，由于农民所面临的不同挑战（主要是病虫害）产量下降了 5%，据一些农民说咖啡的产量更有可能下降，与木薯、香蕉、西红柿等其他作物一样，已经提供了许多对付作物病虫害的方法，但是这些方法是在作物已经受到影响时才可以进行的，在种植前预防农业病害仍然是一个具有挑战性和根本性的问题。

长期以来，有机农业的支持者一直在宣扬这样一种观点，即通过有机农业实践，包括建立和维护健康土壤来降低虫害爆发的可能性。最近的研究表明，植物对病虫害的抗性与土壤的最佳物理、化学和生物学特性有关，在主要农作物中，病虫害和杂草造成相当大的产量损失。

本篇论文的研究，试图不使用杀虫剂来革新病虫害监测程序，通过人工智能收集土壤图像和土壤特性数据来调节土壤病虫害，通过建立一个监测咖啡害虫和疾病的主动监测模型来建立关系条件，帮助咖啡农确定最佳的病虫害管理方法，这将导致咖啡产量的增加。

三、思路方法

本研究使用的数据集由训练集和测试集组成，训练集一共有 4893 张照片，包括 961 张健康的，2230 张 ALS 病（美国叶斑病），1702 张 CS 病（褐斑病）。测试集包括 1209 张照片。

下图是各种土壤的样例：



(a) American Leaf Spot

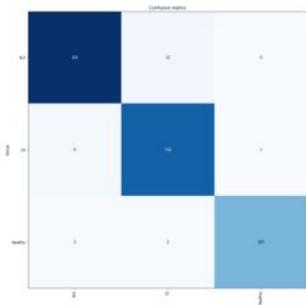


(b) Cercospora Leaf Spot

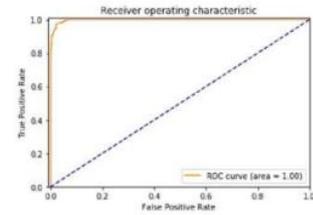


(c) Healthy

残差网络是通过学习特征间的差异引入残差学习，研究者们通过快速学习、转移学习进行了训练，在数据集建模训练时，将训练数据分割成 X 和 Y，以 20% 和 80% 的比例进行了五次训练，结果用混淆矩阵表示。ROC 曲线表示的是混淆矩阵用来衡量模型在各种情况下的性能分类指标，该模型的精确度为 97.3%。



(a) Confusion Matrix.



(b) ROC Curve.

四、实验结论

测试集一共有 1209 张照片，其中 244 张是健康土壤，425 张得 ALS 病，540 张得 CS 病。

研究者在测试数据集上执行了该模型，71%的 ALS 病被正确识别出来，73%的健康样本被正确识别出来，82%的 CS 被正确识别出来，可见识别率比较高。

12.5%的 ALS 被错误识别成了 CS，16.5%的 ALS 被错误识别成了健康，8.2%的 CS 被错误识别成了健康，27%的健康被错误识别成了 ALS，9.8%的 CS 被错误识别成了 ALS。

五、引发思考

通过对上述论文的研究，我对于人工智能的应用有了更加深入的认识。并不局限于我们所了解最多的机器人等，人工智能的应用范围非常之广泛，本文的识别土壤病虫害就是一例。通过大量的机器学习，可以较为准确的提取出土壤的特征，从而判断出土壤的病虫害，让当地的咖啡农可以提前预防，提高产量。人工智能与机器学习发展前景十分广阔，还有更多未知的领域等待探索，随着技术的发展，带给人类更好的生活。